Supporting Information

Heparin remodels the microtubule-binding repeat R3 of Tau protein towards fibril-prone conformations

*Xuewei Dong[a], Ruxi Qi[b], Qin Qiao[c], Xuhua Li[d], Fangying Li[a], Jiaqian Wan[e], Qingwen Zhang[e] and Guanghong Wei[a.]**

a Department of Physics, State Key Laboratory of Surface Physics, and Key Laboratory for Computational Physical Sciences (Ministry of Education), Fudan University, Shanghai 200438, People's Republic of China

b Cryo-EM center, Southern University of Science and Technology, Shenzhen, 518055, People's Republic of China

c Digital Medical Research Center, School of Basic Medical Sciences, Shanghai Key Laboratory of Medical Imaging Computing and Computer Assisted Intervention, Fudan University, Shanghai 200032, People's Republic of China

d MOE Key Laboratory for Nonequilibrium Synthesis and Modulation of Condensed Matter, School of Physics, Xi'an Jiaotong University, Xi'an 710049, People's Republic of China

e College of Physical Education and Training, Shanghai University of Sport, Shanghai 200438, People's Republic of China

Corresponding Author

* E-mail: ghwei@fudan.edu.cn (G. W.)

This supporting information contains convergence assessments of REMD simulations, five supplementary tables (Tables S1-S5) and 13 supplementary figures (Figures S1-S13).

Convergence assessments of REMD simulations

We checked the convergence of REMD simulations by comparing several quantities using data generated within two different time intervals (100-200 and 200-300 ns for mon-R3/mon-R3-Hep systems; 150-225 and 225-300 ns for di-R3/di-R3-Hep systems) at the representative temperature, 310 K. As it can be seen in Figures S4a-c/e-g and Figures S5a-c/e-g, the average probability of each type of secondary structure as well as the probability density function (PDF) of solvent accessible surface area (SASA) and total hydrogen bond (H-bond) number of R3 within the two independent time intervals overlap very well with each other in all the four systems. The time evolution of replica index under 310 K shows that all replicas traveled to 310 K several times during the 300 ns-simulation, indicating that simulations are not trapped in one single temperature (Figures S4d, h and Figures S5d, h). Taken together, these data demonstrate that the REMD simulations of mon-R3/mon-R3-Hep and di-R3/di-R3-Hep systems are reasonably converged within 300 ns.

Five supplementary tables

Table S1. A summary of REMD setup details. For each system, we present system names, size of simulation boxes, simulation time, and the total number of atoms, replicas and initial states.

System	Box size (nm^3)	Simulation time (ns)	Atom #	Replica #	Initial state #
mon-R3	7.6×5.4×5.4 (rectangle)	300	21484	48	1
di-R3	524.19 (dodecahedron)	300	51419	60	6
mon-R3-Hep	335.56 (dodecahedron)	300	33195	54	2
di-R3-Hep	524.19 (dodecahedron)	300	51566	60	6

Table S2. Temperature (K) list used in the 48-replica REMD simulation of R3 monomer without heparin.

310.00	312.33	314.69	317.05	319.42	321.81	324.21	326.62
329.05	331.49	333.95	336.42	338.91	341.41	343.93	346.46
349.02	351.61	354.20	356.77	359.39	361.94	364.58	367.22
369.88	372.60	375.30	378.02	380.76	383.52	386.28	389.07
391.87	394.69	397.53	400.38	403.25	406.14	409.15	411.97
414.91	417.87	420.84	423.84	426.85	429.88	432.93	436.00

Table S3. Temperature (K) list used in the 60-replica REMD simulation of R3 dimer without heparin.

310.00	311.55	313.10	314.66	316.23	317.80	319.37	320.95	322.54	324.14
325.74	327.35	328.97	330.59	332.22	333.85	335.50	337.15	338.80	340.46
342.13	343.80	345.49	347.17	348.86	350.56	352.27	353.98	355.71	357.44
359.17	360.91	362.66	364.41	366.18	367.94	369.72	371.50	373.29	375.09
376.89	378.70	380.52	382.34	384.17	386.01	387.86	389.71	391.57	393.44
395.32	397.20	399.10	401.00	402.89	404.80	406.72	408.65	410.59	412.53

Table S4. Temperature (K) list used in the 54-replica REMD simulation of R3 monomer with heparin.

310.00	311.93	313.87	315.82	317.78	319.75	321.73	323.72	325.72
327.73	329.75	331.77	333.81	335.86	337.92	339.99	342.06	344.14
346.24	348.35	350.47	352.60	354.74	356.89	359.05	361.22	363.40
365.59	367.80	370.01	372.24	374.48	376.72	378.98	381.25	383.55
385.84	388.15	390.46	392.79	395.13	397.48	399.84	402.21	404.60
407.00	409.40	411.82	414.26	416.70	419.16	421.64	424.12	426.62

Table S5. Temperature (K) list used in the 60-replica REMD simulation of R3 dimer with heparin.

310.00	311.55	313.10	314.66	316.23	317.80	319.37	320.95	322.54	324.14
325.74	327.35	328.97	330.59	332.22	333.85	335.50	337.15	338.80	340.46
342.13	343.80	345.49	347.17	348.86	350.56	352.27	353.98	355.71	357.44
359.17	360.91	362.66	364.41	366.18	367.94	369.72	371.50	373.29	375.09
376.89	378.70	380.52	382.34	384.17	386.01	387.86	389.71	391.57	393.44
395.32	397.20	399.10	401.00	402.89	404.80	406.72	408.65	410.59	412.53

13 supplementary figures

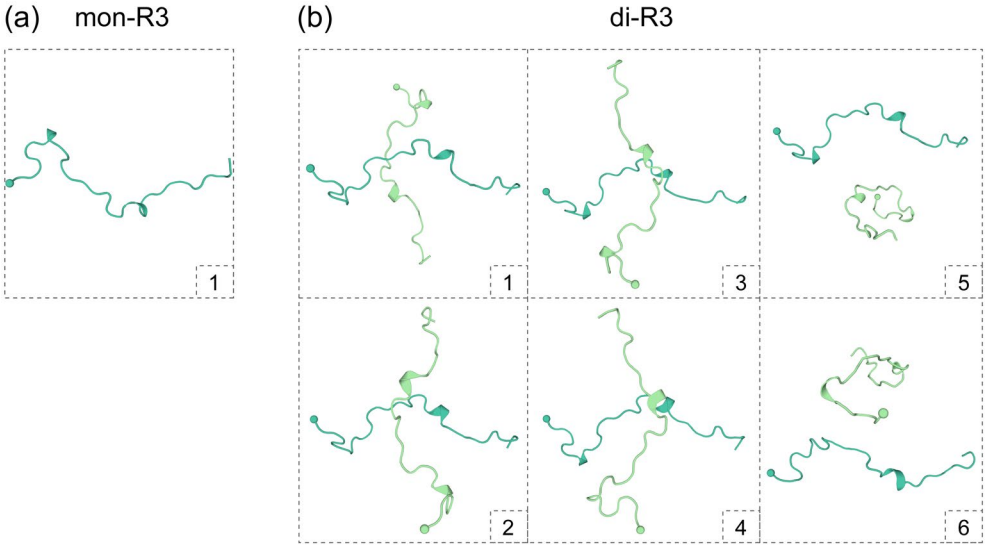

Figure S1. The initial conformations of R3 monomer and dimer in the absence of heparin. R3 is shown in cartoon, with the N-terminal C_α atom of each chain represented by a sphere. For clarity, water molecules and Na^+/Cl^- ions are not shown.

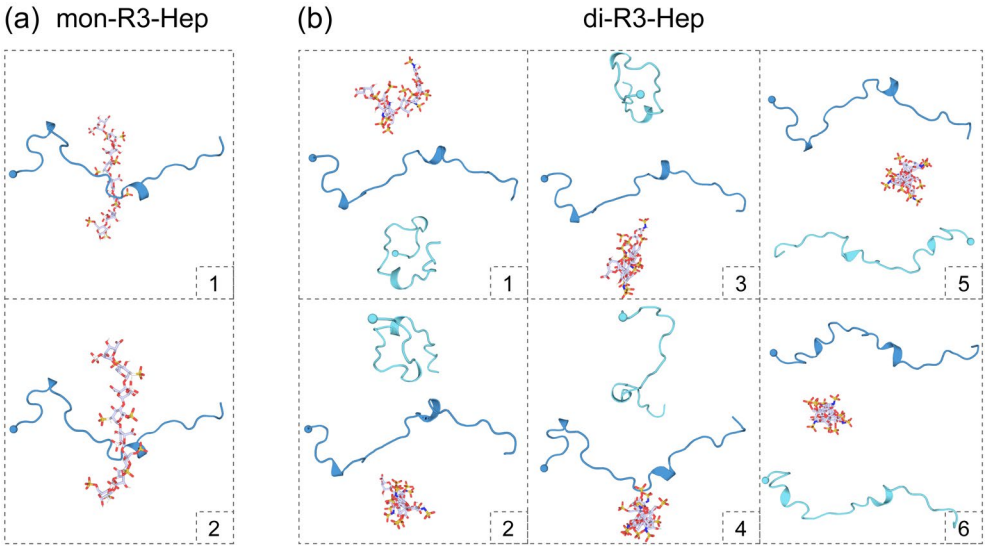

Figure S2. The initial conformations of R3 monomer and dimer in the presence of heparin. R3 is shown in cartoon, with the N-terminal C_α atom of each chain represented by a sphere. Heparin is shown in sticks. The number in each panel is used to label different initial states. For clarity, water molecules and Na^+/Cl^- ions are not shown.

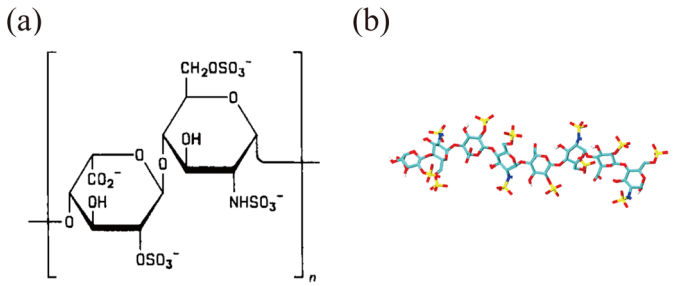

Figure S3. (a) Chemical structure of the IdoA-GlcN-disaccharide unit in heparin. (b) The 3D structure of heparin used in our REMD simulations.

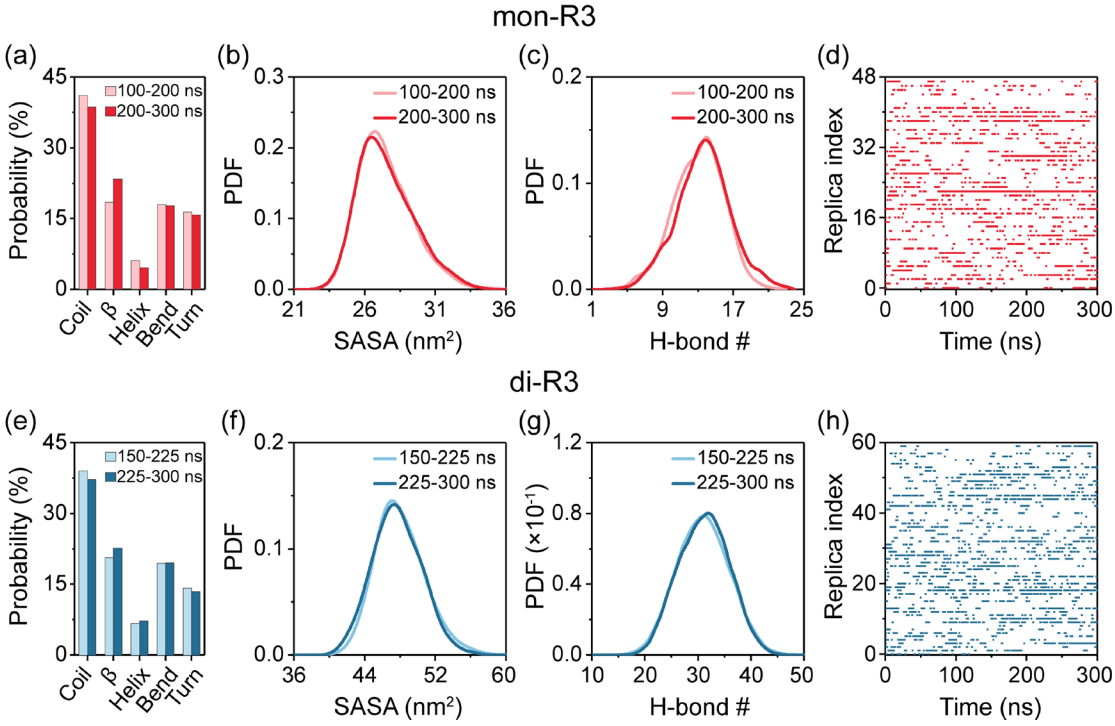

Figure S4. Simulation convergence assessments of mon-R3 and di-R3 systems using data generated within two different time intervals from the first (310 K) temperature replica (100-200 and 200-300 ns for mon-R3 system; 150-225 and 225-300 ns for di-R3 system). (a) The average probability of each type of secondary structure. The PDF of (b) SASA and (c) H-bond number of R3. (d) The time evolution of replica index under the representative temperature of 310 K.

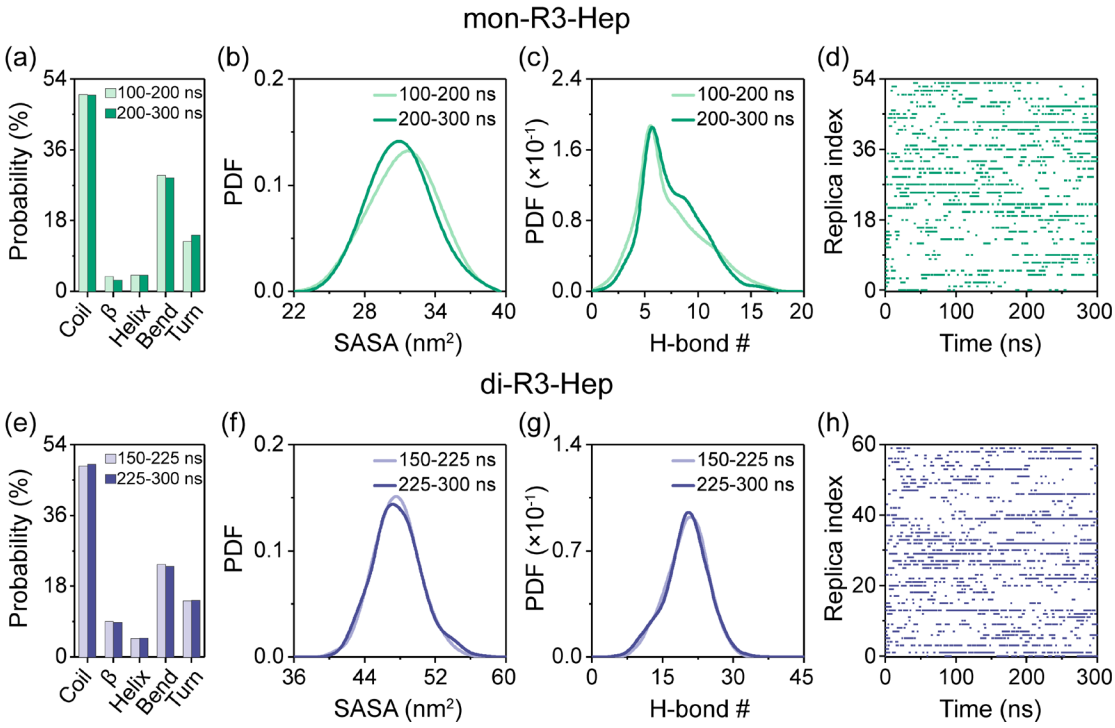

Figure S5. Simulation convergence assessments of mon-R3-Hep and di-R3-Hep systems using data generated within two different time intervals from the first (310 K) temperature replica (100-200 and 200-300 ns for mon-R3-Hep system; 150-225 and 225-300 ns for di-R3-Hep system). (a) The average probability of each type of secondary structure. The PDF of (b) SASA and (c) H-bond number of R3. (d) The time evolution of replica index under the representative temperature of 310 K.

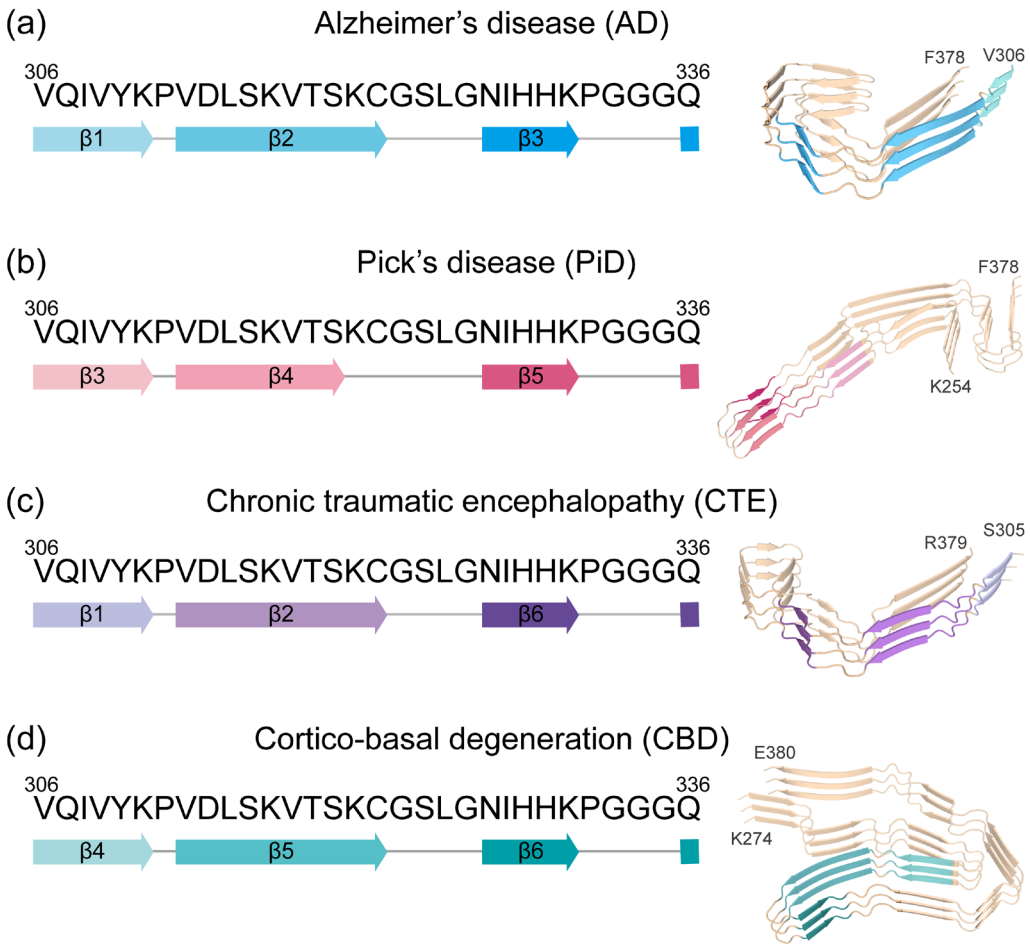

Figure S6. Disease-specific Tau protofilament cores. Amino acid sequence and β-strand regions of R3 in Tau filament cores and structures of Tau protofilaments from patients with different tauopathies: (a) Alzheimer's disease (AD, PDB ID: 5O3L),[1] (b) Pick's disease (PiD, PDB ID: 6GX5),[2] (c) chronic traumatic encephalopathy (CTE, PDB ID: 6NWP)[3] and (d) cortico-basal degeneration (CBD, PDB ID: 6VHA).[4]

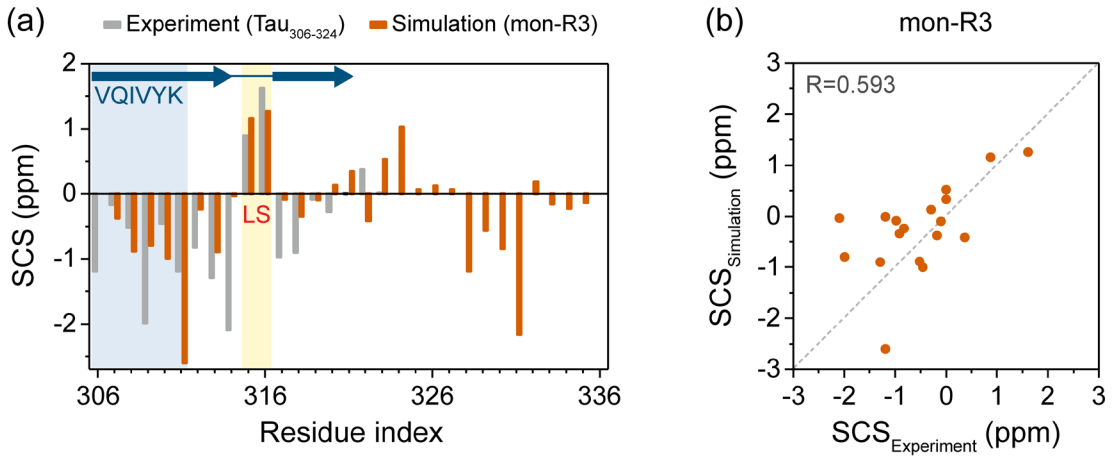

Figure S7. Secondary chemical shift (SCS) analysis of R3. (a) Comparison of C_α SCSs of R3 between experimental (gray) and simulation-predicted (orange) values. The β and turn regions are highlighted by arrows and line. ^{306}VQIVYK^{311} and ^{315}LS^{316} are highlighted in blue and yellow shadows. (b) Scatter plots show the comparison between $SCS_{Experiment}$ and $SCS_{Simulation}$ of C_α atoms for R3, with the Pearson correlation coefficient (R) = 0.593.

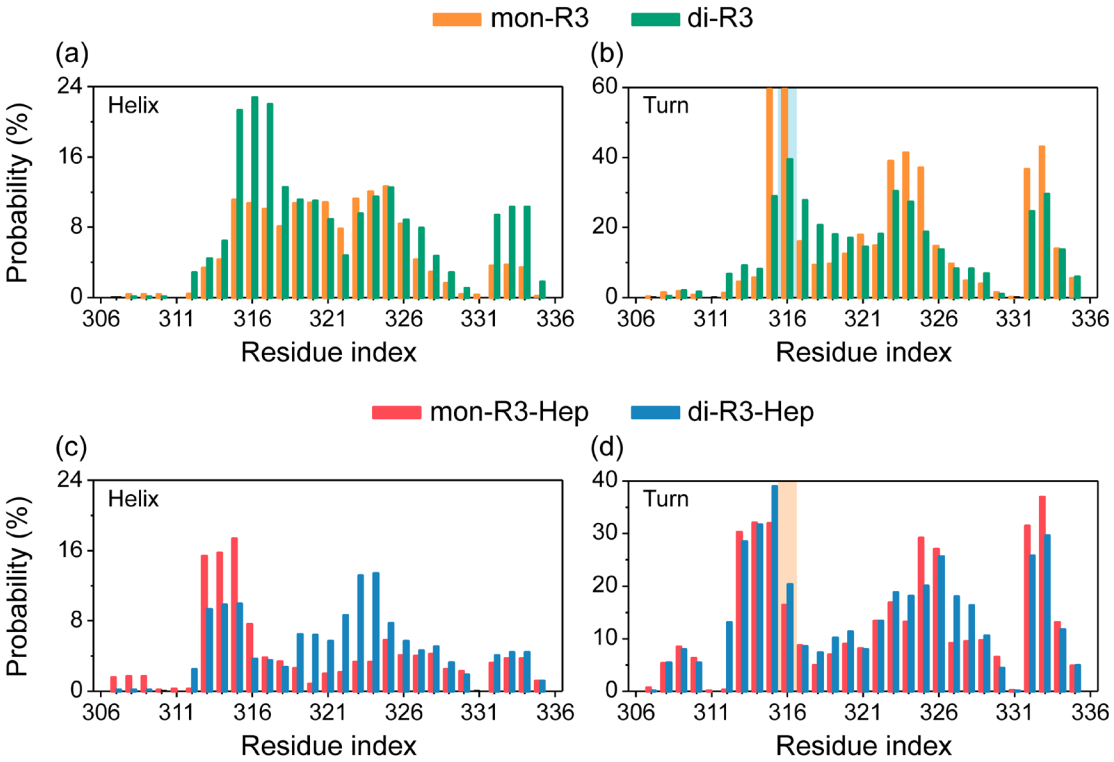

Figure S8. Residue-based (a, c) helix and (b, d) turn probabilities of R3 monomer and dimer in the absence (top panel) and presence (bottom panel) of heparin. Residue S316 is highlighted in blue or orange shadow.

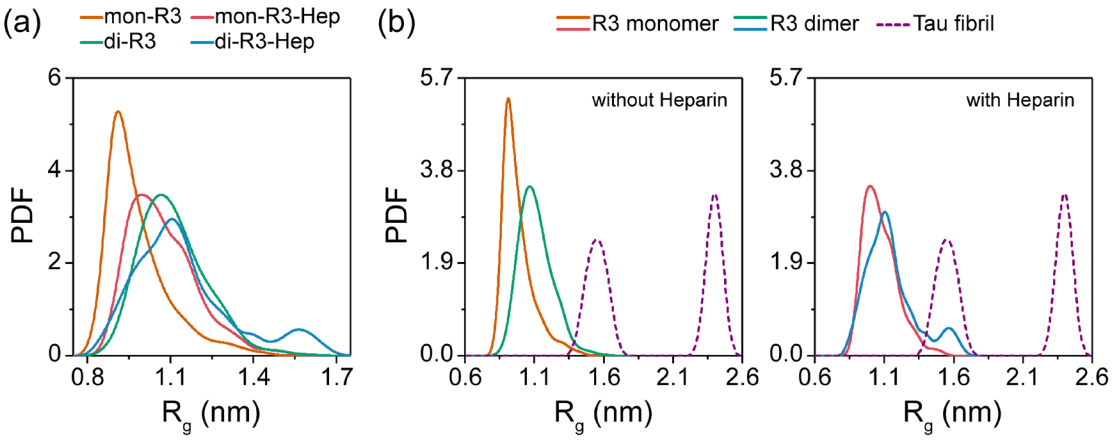

Figure S9. The PDF of R_g of (a) R3 monomer and dimer with and without heparin, compared to that of (b) R3 segment in the four disease-specific cryo-EM-derived filaments (PDB ID: 5O3L for AD, 6GX5 for PiD, 6NWP for CTE and 6VHA for CBD).

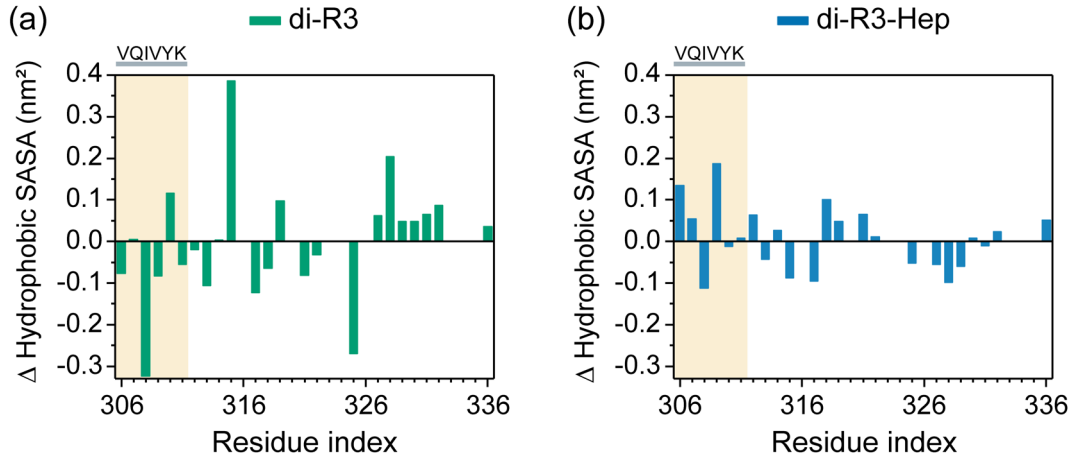

Figure S10. Differences between residue-based hydrophobic SASAs of conformations of R3 with a single-chain R_g > 1.5 nm and those averaged over all conformations in (a) di-R3 and (b) di-R3-Hep systems. The fibril-nucleating core, PHF6 (^{306}VQIVYK^{311}), is highlighted in yellow shadow.

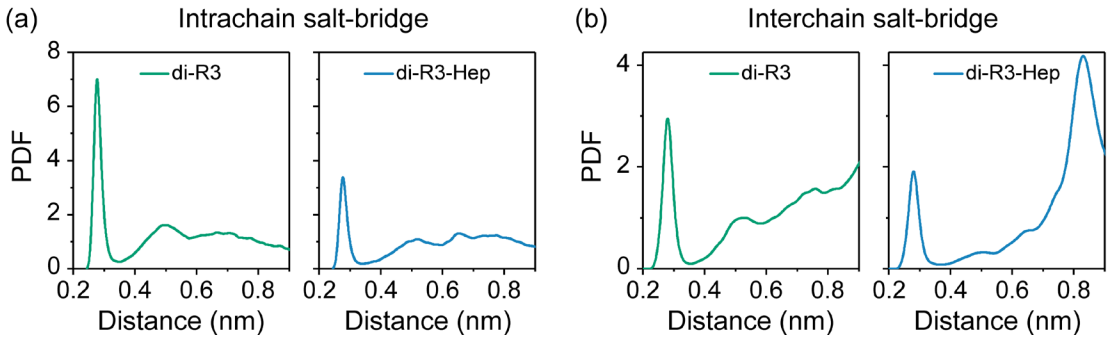

Figure S11. Effect of heparin on intrachain and interchain salt bridges. The PDFs of the minimum distance between (a) intrachain/(b) interchain Nζ atoms of the side-chain NH_3^+ groups of all positively charged Lys residues and the COO^- group of negatively charged Asp in di-R3 and di-R3-Hep systems.

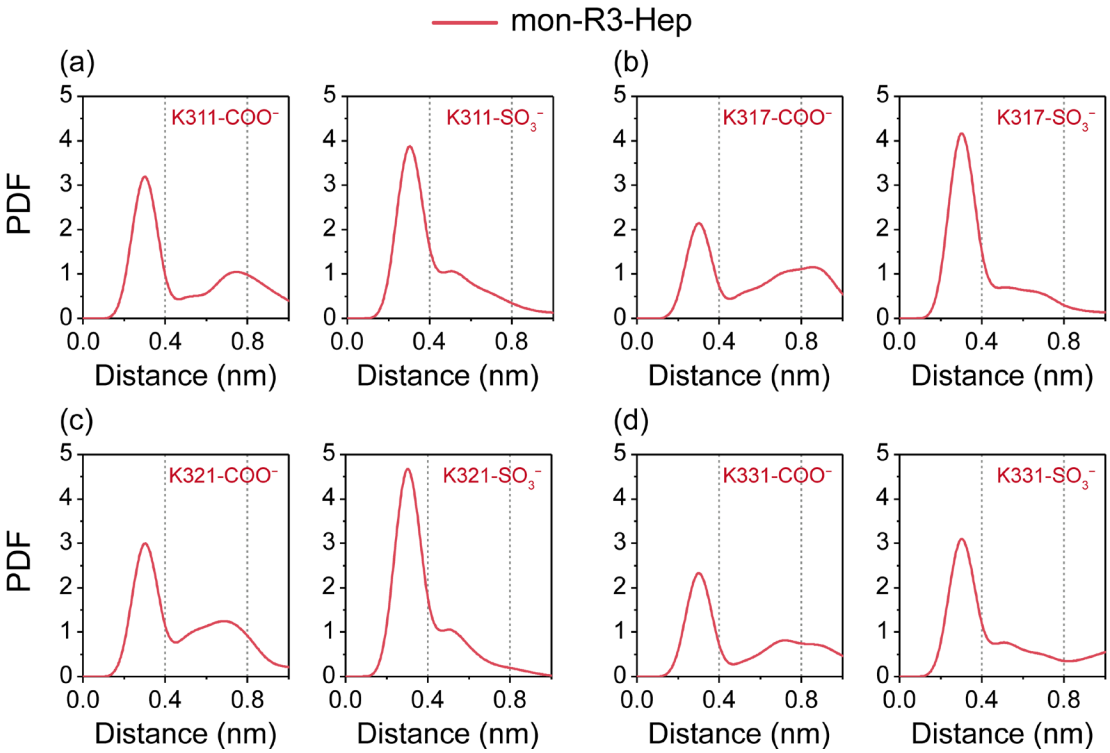

Figure S12. Analysis of salt bridge interactions between R3 monomer and heparin. PDFs of the minimum distance between Nζ atom of the sidechain NH_3^+ group of (a) K311/(b) K317/(c) K321/(d) K331 and the COO^-/SO_3^- groups of heparin in mon-R3-Hep system.

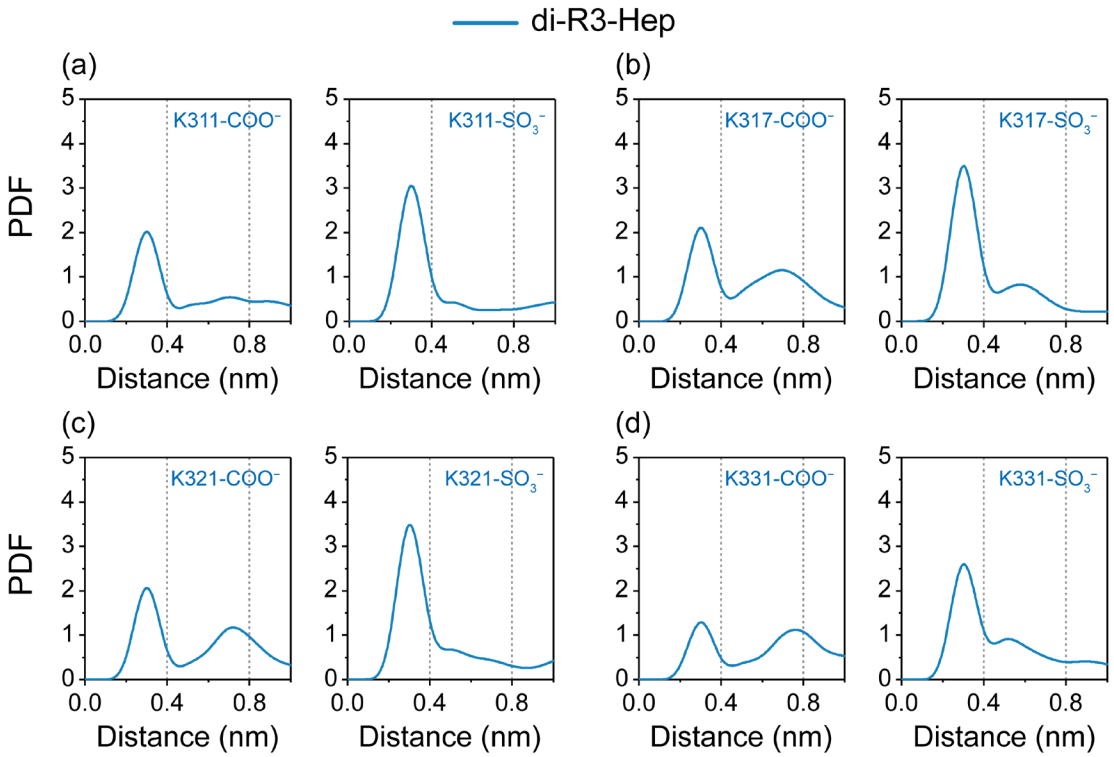

Figure S13. Analysis of salt bridge interactions between R3 dimer and heparin. PDFs of the minimum distance between Nζ atoms of the sidechain NH_3^+ groups of (a) K311/(b) K317/(c) K321/(d) K331 and the COO^-/SO_3^- groups of heparin in di-R3-Hep system.

REFERENCES

1. A. W. P. Fitzpatrick, B. Falcon, S. He, A. G. Murzin, G. Murshudov, H. J. Garringer, R. A. Crowther, B. Ghetti, M. Goedert and S. H. W. Scheres, *Nature*, 2017, **547**, 185-190.
2. B. Falcon, W. Zhang, A. G. Murzin, G. Murshudov, H. J. Garringer, R. Vidal, R. A. Crowther, B. Ghetti, S. H. W. Scheres and M. Goedert, *Nature*, 2018, **561**, 137-140.
3. B. Falcon, J. Zivanov, W. Zhang, A. G. Murzin, H. J. Garringer, R. Vidal, R. A. Crowther, K. L. Newell, B. Ghetti, M. Goedert and S. H. W. Scheres, *Nature*, 2019, **568**, 420-423.
4. T. Arakhamia, C. E. Lee, Y. Carlomagno, D. M. Duong, S. R. Kundinger, K. Wang, D. Williams, M. DeTure, D. W. Dickson, C. N. Cook, N. T. Seyfried, L. Petrucelli and A. W. P. Fitzpatrick, *Cell*, 2020, **180**, 633-644.e12.